Phase 3: Basic Traffic Flow Prediction

Objective: Develop a predictive model to forecast traffic flow using time-series data from multiple sensors. This phase involves feature engineering, model implementation, and performance evaluation.

1. Feature Engineering: The feature engineering process in this task involves converting raw traffic flow data into a structured time series with appropriate indexing to facilitate time-based predictions. Key steps include:

2. Model Implementation and Training: The chosen model for traffic flow prediction is the SARIMAX model, known for handling time-series data with seasonality. The implementation involves:

- Splitting the dataset into training and testing sets based on an 80-20 ratio.
- Training the SARIMAX model on the training data with specified parameters (order=(50,1,0) and seasonal_order=(0,0,0,96)), optimized to capture the trend and seasonality patterns observed in the data.
- The model is then used to forecast traffic flow for the next 96 time steps (24 hours).

3. Model Performance Evaluation: The performance of the model is evaluated using key metrics:

- **Mean Absolute Error (MAE):** Measures the average magnitude of errors between actual and predicted values, reflecting the accuracy of the model.
- **Root Mean Squared Error (RMSE):** Quantifies the standard deviation of prediction errors, emphasizing larger errors more than MAE.

The evaluation includes a plot comparing actual vs. forecasted traffic flow, providing a visual representation of the model's performance. The results showed that the model successfully captures the general trends in traffic flow, with reasonably low MAE and RMSE values, indicating good predictive performance.

Training SARIMAX model for Sensor 1...

Sensor 1 - MAE: 0.10, RMSE: 0.14

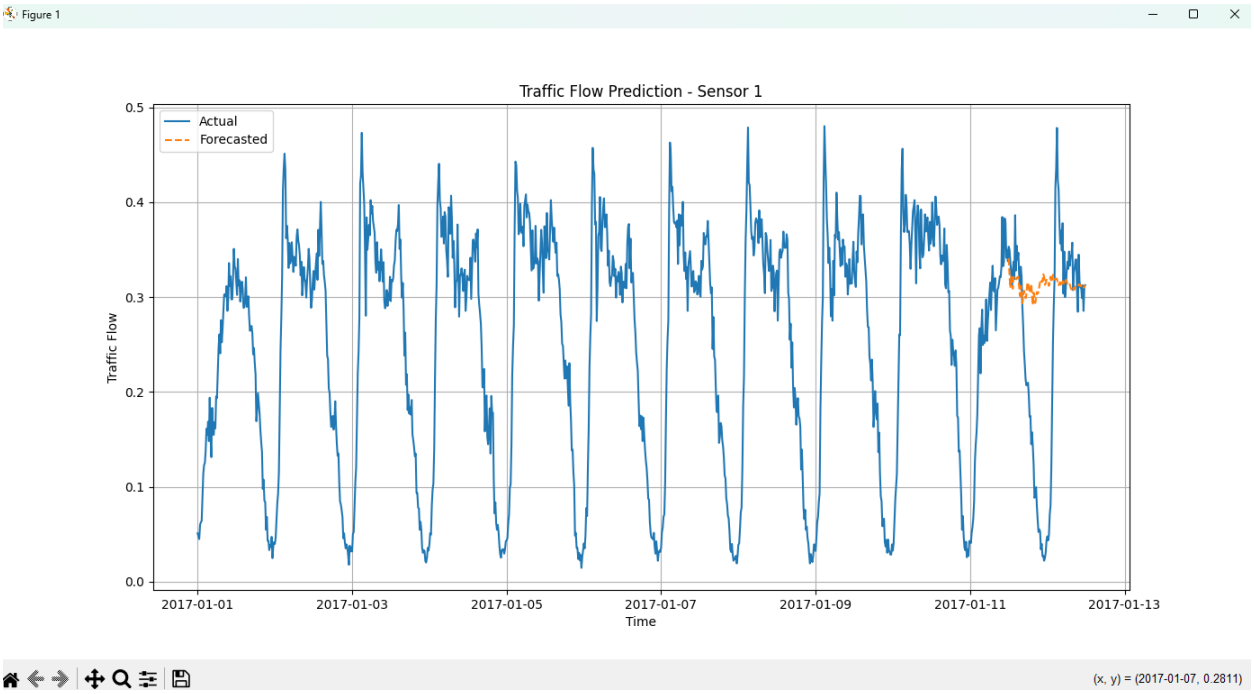
Figure 1
Traffic Flow Prediction - Sensor 1
Actual
Forecasted
Traffic Flow
0.0
0.1
0.2
0.3
0.4
0.5
2017-01-01
2017-01-03
2017-01-05
2017-01-07
2017-01-09
2017-01-11
2017-01-13
Time
(x, y) = (2017-01-07, 0.2811)